

# Isogeometric Least-squares Collocation Method with Consistency and Convergence Analysis

Hongwei Lin<sup>a,\*</sup>, Yunyang Xiong<sup>a</sup>, Qianqian Hu<sup>b</sup>

<sup>a</sup>*Department of Mathematics, State Key Lab. of CAD&CG, Zhejiang University, Hangzhou, 310027, China*

<sup>b</sup>*Department of Mathematics, Zhejiang Gongshang University, Hangzhou, 310018, China*

---

## Abstract

Isogeometric analysis (IGA) approximates the unknown solution of a boundary value problem (or initial value problem) using a non-uniform rational basis spline (NURBS) function with non-linear NURBS basis functions. If the order of the NURBS function is high enough, the boundary value problem can be solved using an isogeometric collocation (IGA-C) method that interpolates the real differential operator at collocation points. In this paper, we present the isogeometric least-squares collocation (IGA-L) method. IGA-L determines the numerical solution by making the approximate differential operator fit the real differential operator in a least-squares sense. Therefore, while IGA-C requires the number of collocation points to be equal to that of the unknown coefficients of the numerical solution, the number of collocation points employed in IGA-L can be larger than that of the unknown coefficients. IGA-L has some advantages over IGA-C. First, if the number of coefficients of the numerical solution is fixed, a small increase in the number of collocation points in IGA-L leads to a large improvement in the accuracy of its numerical solution compared to IGA-C. Second, because the number of collocation points in IGA-L is variable, it is more flexible and more stable than IGA-C. We also show the consistency and convergence properties of IGA-L in this paper.

**Keywords:** Isogeometric analysis, collocation method, least-squares fitting, NURBS, consistency and convergence

---

## 1. Introduction

While classical Finite Element Analysis (FEA) methods, widely employed in physical simulations, are based on element-wise piecewise polynomials, computer-aided design (CAD) models are usually represented by non-uniform rational basis splines (NURBSs) with non-linear NURBS basis functions. Therefore, the first task in a CAD model simulation is to transform the non-linear NURBS-represented CAD model into a linear mesh representation. This mesh transformation is a very tedious operation, and it has become the most time-consuming task of the whole FEA procedure. To avoid the mesh transformation and advance the seamless integration of CAD and computer-aided engineering (CAE), isogeometric analysis (IGA) was invented

---

\*Corresponding author: phone number: 86-571-87951860-8304, fax number: 86-571-88206681, email: hwlin@zju.edu.cn

Preprint submitted to Elsevier

January 28, 2016

by Hughes et al. [1]. IGA is based on the non-linear NURBS basis functions, hence it can deal with NURBS-represented CAD models directly. In this way, IGA not only saves a significant amount of computation, it also greatly improves the numerical accuracy of the solution.

In IGA, the analytical solution  $T$  to a boundary value problem is approximated by a NURBS function  $T_r$  with unknown coefficients. (For brevity, we only mention the boundary value problem in this paper. However, this method is also suitable for the initial value problem.) Solving the boundary value problem is then equivalent to determining the unknown coefficients of  $T_r$ . If the order of  $T_r$  is higher than that of the differential operator  $\mathcal{D}$  of the boundary value problem,  $\mathcal{D}T_r$  can be represented explicitly by a NURBS derivative formula. Therefore, collocation methods can be applied to the strong form of the boundary value problem to determine the unknown coefficients of  $T_r$ .

In Ref. [2], an isogeometric collocation (IGA-C) method was presented. Suppose  $n$  is the number of unknown coefficients of  $T_r$ . IGA-C first samples  $n$  values  $\mathcal{D}T(\eta_j)$ ,  $j = 1, 2, \dots, n$ , and then generates a linear system of equations by making  $\mathcal{D}T_r$  interpolate the  $n$  values, i.e.,  $\mathcal{D}T_r(\eta_j) = \mathcal{D}T(\eta_j)$ ,  $j = 1, 2, \dots, n$ . In this way, the unknown coefficients of  $T_r$  can be determined by solving the linear system.

Essentially, IGA-C acquires the unknown coefficients by interpolation, where the number of the collocation points must be equal to that of the unknown coefficients. In this paper, we propose the isogeometric least-squares collocation (IGA-L) method, which allows the number of collocation points to be larger than that of the unknown coefficients. It yields some advantages over IGA-C. Instead of interpolation, IGA-L makes use of approximation to determine the unknown coefficients of  $T_r$ . Specifically, IGA-L first samples  $m$  values  $\mathcal{D}T(\eta_j)$ ,  $j = 1, 2, \dots, m$ , where  $m$  is greater than the number of unknown coefficients, i.e.,  $m > n$ . The coefficients of the unknown solution  $T_r$  are then determined by solving the least-squares fitting problem

$$\min \sum_{j=1}^m \|\mathcal{D}T(\eta_j) - \mathcal{D}T_r(\eta_j)\|^2.$$

There are two advantages of IGA-L over IGA-C. First, numerical tests presented in this paper show that a small increase in the computation of IGA-L leads to a large improvement in the numerical accuracy of the solution, even though the computational cost of IGA-L is only slightly more than that of IGA-C. Second, IGA-L is more flexible and more stable than IGA-C, because IGA-L allows a variable number of collocation points, while the number of collocation points in IGA-C is fixed to be equal to the number of control points.

The structure of this paper is as follows. In Section 1.1, some related work is briefly reviewed. In Section 2, the generic formulation of IGA-L is presented. In Section 3, we show the consistency and convergence properties of IGA-L. After thoroughly comparing IGA-L and IGA-C, both in theory and with numerical examples in Section 4, this paper is concluded in Section 5.

### 1.1. Related Work

**Least-squares collocation methods:** Although collocation-based meshless methods are efficient, equilibrium conditions are satisfied only at collocation points. Thus, collocation-based meshless methods can result in significant error. To improve computational accuracy, Zhang et al. developed a least-squares collocation method [3], where equilibrium conditions hold at both collocation points and auxiliary points in a least-squares sense. In order to generate a better conditioned linear algebraic equations, a least-squares meshfree collocation method was proposed, based on first-order differential equations [4]. In Ref. [5], a point collocation method

was invented that employs the approximating derivatives based on the moving least-square reproducing kernel approximations. Moreover, several schemes using least-squares collocation methods were developed for two- and three-dimensional heat conduction problems [6], transient and steady-state hyperbolic problems [7], and adaptive analysis problems in linear elasticity [8].

However, the consistency and convergence properties of the aforementioned least-squares collocation methods were only validated using numerical examples, and theoretical numerical analyses were not reported. In this paper, we not only develop an isogeometric least-squares collocation method, but also show its consistency and convergence properties in theory.

**Isogeometric analysis:** NURBS is a mathematical model for representing curves and surfaces by blending weighted control points and NURBS basis functions. Hence, the shape of the NURBS curves and surfaces can be easily modified by moving their control points. Because of the desirable traits of NURBS basis functions, NURBSs curves and surfaces have many good properties, such as convex hull, affine invariance, and variation diminishing. Moreover, NURBSs can represent conic sections and quadric surfaces accurately. Therefore, NURBSs have been widely used in CAD, computer-aided manufacturing (CAM), and CAE, and have become a part of numerous industry standards, such as IGES, STEP, ACIS, and PHIGS. For more details on NURBSs, there are excellent books on the subject [9, 10].

While a NURBS employs non-linear basis functions, classical FEA is based on element-wise piecewise polynomials. Hence, when analyzing NURBS-represented CAD models using classical FEA, the CAD model should be discretized into a mesh model. The mesh generation procedure not only yields an approximation, it is also very tedious, and hence has become a bottleneck in FEA. To overcome these difficulties, Hughes et al. invented the IGA technique [1]. Because IGA is based on NURBS basis functions, it can handle NURBS-represented CAD models directly without generating a mesh. Moreover, because NURBSs can represent complex shapes (and physical fields) with significantly fewer control points than a mesh representation, the computational efficiency and numerical accuracy of IGA are higher than the classical FEA method. In addition, because of the knot insertion property of NURBSs, the original shape of the CAD model can be maintained exactly in the refinement procedure [1].

In geometric design, the NURBS representation is usually employed to model curves and surfaces. There are a few effective methods in geometric design for modeling spline solids. To model a NURBS solid for IGA applications, Zhang et al. proposed a solid construction method from a boundary representation [11]. In the study of IGA, making the geometric representation more suitable for analysis purposes is a key research goal. Cohen et al. presented the analysis-aware modeling technique [12]. Moreover, T-spline [13, 14], trimmed surfaces [15], and subdivision solids [16] were also employed in the IGA method to model the physical domain.

Currently, the IGA method has been successfully applied in various simulation problems, including elasticity [17, 18], structure [19, 20, 21], and fluids [22, 23, 24]. On the other hand, some work concerns the computational aspect of the IGA method, for instance, to improve the accuracy and efficiency by reparameterization and refinement [25, 26, 27, 28, 29].

**Isogeometric collocation methods:** The collocation method is a simple numerical method for solving differential equations that can generate a solution that satisfies the differential equation at a set of discrete points, called collocation points [30]. If the order of the unknown NURBS function that is employed to approximate the solution of the differential equation is high enough, the collocation method can be applied to the strong form of the differential equation. In this way, the IGA-C method was presented in Ref. [2]. However, IGA-C is greatly influenced by the collocation points. Recently, a comprehensive study on IGA-C discovered its superior behavior over the Galerkin method in terms of its accuracy-to-computational-time ratio [31], and the

consistency and convergence properties of the IGA-C method were developed in Ref. [32].

Moreover, based on the local hierarchical refinement of NURBSs, adaptive IGA-Cs were developed and analyzed [31]. Meanwhile, IGA-C has also been extended to multi-patch NURBS configurations, various boundary and patch interface conditions, and explicit dynamic analysis [33]. Recently, IGA-C was successfully employed to solve the Timoshenko beam problem [34] and spatial Timoshenko rod problem [35], showing that mixed collocation schemes are locking-free, independently of the choice of the polynomial degrees for the unknown fields.

## 2. Generic Formulation of IGA-L

Consider the following boundary value problem,

$$\begin{cases} \mathcal{D}T = f, & \text{in } \Omega \in \mathbb{R}^d, \\ \mathcal{G}T = g, & \text{on } \partial\Omega, \end{cases} \quad (1)$$

where  $\Omega$  is the physical domain in  $\mathbb{R}^d$ ,  $\mathcal{D}$  is a differential operator in the physical domain,  $\mathcal{G}T = g$  is the boundary condition, and  $f : \Omega \rightarrow \mathbb{R}$  and  $g : \partial\Omega \rightarrow \mathbb{R}$  are given functions. Suppose  $d_1$  is the maximum order of derivatives appearing in  $\mathcal{D} : V \rightarrow W$ , where  $V$  and  $W$  are two Hilbert spaces, and the analytical solution  $T \in C^{d_2}(\Omega)$ ,  $d_2 > d_1 \geq 1$ .

In IGA, the physical domain  $\Omega$  is represented by a NURBS mapping:

$$\mathbf{F} : \Omega_0 \rightarrow \Omega, \quad (2)$$

where  $\Omega_0$  is the parametric domain. By replacing the control points of  $\mathbf{F}(\Omega_0)$  with unknown control coefficients, the representation of the unknown numerical solution  $T_r$  is generated.

Suppose there are  $n$  unknown control coefficients in the unknown numerical solution  $T_r$ . We first sample  $m_1$  points  $\theta_i$  inside the parametric domain  $\Omega_0$  that correspond to  $m_1$  values  $\eta_i = \mathbf{F}(\theta_i)$ ,  $i = 1, 2, \dots, m_1$ , inside the physical domain  $\Omega$ . Furthermore, we sample  $m_2$  points  $\theta_j$  on the parametric domain boundary  $\partial\Omega_0$  that correspond to  $m_2$  values  $\eta_j = \mathbf{F}(\theta_j)$ ,  $j = m_1 + 1, m_1 + 2, \dots, m_1 + m_2$ , on the physical domain boundary  $\partial\Omega$ . The total number of these points, i.e.,  $m = m_1 + m_2$ , is greater than the number of unknown coefficients of the numerical solution  $T_r$ , namely,  $m = m_1 + m_2 > n$ . Just as in IGA-C, these sampling points are also called **collocation points**.

Substituting these collocation points into the boundary value problem (1), a system of equations with  $m$  equations and  $n$  unknowns is obtained (where  $m = m_1 + m_2 > n$ ),

$$\begin{cases} \mathcal{D}T_r(\eta_i) = f(\eta_i), & i = 1, 2, \dots, m_1, \\ \mathcal{G}T_r(\eta_j) = g(\eta_j), & j = m_1 + 1, m_1 + 2, \dots, m_1 + m_2. \end{cases} \quad (3)$$

Arranging the unknowns of the numerical solution  $T_r$  into an  $n \times 1$  matrix, i.e.,  $X = [x_1 \ x_2 \ \dots \ x_n]^t$ , the system of equations (3) can be represented in matrix form by

$$AX = b.$$

Because the number of equations is greater than the number of unknowns, the solution is sought in the least-squares sense, i.e.,

$$\min_X \|AX - b\|^2. \quad (4)$$

**Computation of the least-squares problem (4):** The least-squares problem (4) is very important in practice, and there have been developed lots of robust and efficient methods for solving it [36]. One frequently employed method is to solve the normal equation (5),

$$A^T A X = A^T b. \quad (5)$$

Although the condition number of the matrix  $A^T A$  is the square of that of the matrix  $A$ ,  $A^T A$  is a symmetric positive definite matrix, and the normal equation (5) can be solved efficiently by Cholesky decomposition [36]. Moreover, Householder orthogonalization and Given orthogonalization are also two efficient methods [36] which often employed in solving the least-squares problem (4). For more methods on solving (4), please refer to [36].

In the following, we will show the consistency and convergence properties of IGA-L, and compare it with IGA-C, both in theory and with numerical examples.

### 3. Numerical Analysis

In the IGA-L method developed above, a NURBS function  $T_r$  is employed to approximate the analytical solution  $T$  of the boundary value problem (1), and hence the real differential operator  $\mathcal{D}T$  is approximated by  $\mathcal{D}T_r$ . In Ref. [32], it was shown that both  $\mathcal{D}T_r$  and  $T_r$  are defined on the same knot intervals, i.e.,

**Lemma 1.** *If  $T_r$  is a NURBS function and  $\mathcal{D}$  is a differential operator, then both  $\mathcal{D}T_r$  and  $T_r$  are defined on the same knot intervals [32].*

Moreover, given a set  $\Phi$ , the **diameter** of  $\Phi$  is defined as

$$\text{diam}(\Phi) = \max\{d(x, y), x, y \in \Phi\},$$

where  $d(x, y)$  is the Euclidean distance between  $x$  and  $y$ . In the following, we suppose the NURBS function  $T_r$  to be defined on a knot grid  $\mathcal{T}^h$ , where  $h$  is the **knot grid size**: In the one-dimensional case,  $\mathcal{T}^h$  is a knot sequence, and  $h = \max_i\{\text{diam}([u_i, u_{i+1}])\}$ ; in the two-dimensional case,  $\mathcal{T}^h$  is a rectangular grid, and  $h = \max_{ij}\{\text{diam}([u_i, u_{i+1}] \times [v_j, v_{j+1}])\}$ ; and in the three-dimensional case,  $\mathcal{T}^h$  is a hexahedral grid, and  $h = \max_{ijk}\{\text{diam}([u_i, u_{i+1}] \times [v_j, v_{j+1}] \times [w_k, w_{k+1}])\}$ .

In this section, we study the consistency and convergence properties of IGA-L, i.e., when  $h \rightarrow 0$ , not only will approximate differential operator  $\mathcal{D}T_r$  tend to real operator  $\mathcal{D}T$ , but numerical solution  $T_r$  will also tend to analytical solution  $T$ .

#### 3.1. Consistency

In this section, we explore the consistency property of the IGA-L method. The theorem for consistency is presented as follows.

**Theorem 1.** *In the IGA-L method, if*

- (1.1) *each knot interval of the NURBS function  $T_r$  defined on knot grid  $\mathcal{T}^h$  contains at least one collocation point*
- (1.2) *the degree of each variable in  $T_r$  is larger than the maximum order of the partial derivatives to the variables appearing in  $\mathcal{D}$  (see Eq. (1))*

(1.3) the least-squares fitting error  $e_h$  is bounded, i.e.,  $e_h < \bar{M}$ , where  $\bar{M}$  is a positive constant

then approximate differential operator  $\mathcal{DT}_r$  tends to real differential operator  $\mathcal{DT}$  in the  $L^2$  norm when  $h \rightarrow 0$ .

**Proof:** We only show the theorem in the two-dimensional case. The proof for the one- and three-dimensional cases is similar.

In the two-dimensional case, suppose the tensor product NURBS function  $T_r(u, v)$  of degree  $l_u \times l_v$  is defined on the knot sequences

$$\begin{aligned} & \underbrace{\{u_0, u_0, \dots, u_0, u_1, \dots, u_{n_u-1}, u_{n_u}, u_{n_u}, \dots, u_{n_u}\}}_{l_u+1}, \\ & \underbrace{\{v_0, v_0, \dots, v_0, v_1, \dots, v_{n_v-1}, u_{n_v}, u_{n_v}, \dots, u_{n_v}\}}_{l_v+1}. \end{aligned} \quad (6)$$

Then the corresponding knot grid is

$$\mathcal{T}^h = \{[u_i, u_{i+1}] \times [v_j, v_{j+1}], i = 0, 1, \dots, n_u - 1, j = 0, 1, \dots, n_v - 1\}. \quad (7)$$

Denote

$$R(u, v) = (\mathcal{DT}(u, v) - \mathcal{DT}_r(u, v))^2, (u, v) \in [u_0, u_{n_u}] \times [v_0, v_{n_v}].$$

Note that  $T_r(u, v)$  is generated by least-squares fitting the values of  $\mathcal{DT}(u, v)$  at the collocation points  $\boldsymbol{\vartheta}_d = (\eta_d, \xi_d)$ , i.e.,  $\mathcal{DT}(\boldsymbol{\vartheta}_d)$ ,  $d = 1, 2, \dots, D$ , ( $D \geq n_u n_v$ ). And suppose the fitting error is

$$e_h = \sum_{k=1}^D R(\boldsymbol{\vartheta}_d) = \sum_{k=1}^D (\mathcal{DT}(\boldsymbol{\vartheta}_d) - \mathcal{DT}_r(\boldsymbol{\vartheta}_d))^2. \quad (8)$$

First, based on Lemma 1, the numerical solution  $T_r(u, v)$  and the approximate differential operator  $\mathcal{DT}_r(u, v)$  have the same knot intervals,

$$[u_i, u_{i+1}] \times [v_j, v_{j+1}], i = 0, 1, \dots, n_u - 1, j = 0, 1, \dots, n_v - 1.$$

Now, consider the error between  $\mathcal{DT}(u, v)$  and  $\mathcal{DT}_r(u, v)$  in the  $L^2$  norm,

$$\|\mathcal{DT}(u, v) - \mathcal{DT}_r(u, v)\|_{L^2}^2 = \int_{v_0}^{v_{n_v}} \int_{u_0}^{u_{n_u}} R(u, v) du dv = \sum_{j=0}^{n_v-1} \sum_{i=0}^{n_u-1} \int_{v_j}^{v_{j+1}} \int_{u_i}^{u_{i+1}} R(u, v) du dv.$$

Since each knot interval  $[u_i, u_{i+1}] \times [v_j, v_{j+1}]$  contains at least one collocation point, we suppose  $\boldsymbol{\vartheta}_d = (\eta_d, \xi_d) \in [u_i, u_{i+1}] \times [v_j, v_{j+1}]$ . Using the left and right rectangle integral formula repeatedly,

we get

$$\begin{aligned}
\int_{v_j}^{v_{j+1}} \int_{u_i}^{u_{i+1}} R(u, v) du dv &= \int_{v_j}^{v_{j+1}} dv \left( \int_{u_i}^{\eta_d} R(u, v) du + \int_{\eta_d}^{u_{i+1}} R(u, v) du \right) \\
&= \int_{v_j}^{v_{j+1}} \left( (\eta_d - u_i) R(\eta_d, v) + (u_{i+1} - \eta_d) R(\eta_d, v) + (\eta_d - u_i)^2 \frac{R'_u(\mu_i^{(1)}(v), v)}{2} + (u_{i+1} - \eta_d)^2 \frac{R'_u(\mu_i^{(2)}(v), v)}{2} \right) dv \\
&= (u_{i+1} - u_i) \int_{v_j}^{v_{j+1}} R(\eta_d, v) dv + (\eta_d - u_i)^2 \int_{v_j}^{v_{j+1}} \frac{R'_u(\mu_i^{(1)}(v), v)}{2} dv + (u_{i+1} - \eta_d)^2 \int_{v_j}^{v_{j+1}} \frac{R'_u(\mu_i^{(2)}(v), v)}{2} dv \\
&= (u_{i+1} - u_i) \left( (\xi_d - v_j) R(\eta_d, \xi_d) + (v_{j+1} - \xi_d) R(\eta_d, \xi_d) + (\xi_d - v_j)^2 \frac{R'_v(\eta_d, \omega_{ij}^{(1)})}{2} + (v_{j+1} - \xi_d)^2 \frac{R'_v(\eta_d, \omega_{ij}^{(2)})}{2} \right) \\
&\quad + (\eta_d - u_i)^2 \int_{v_j}^{v_{j+1}} \frac{R'_u(\mu_i^{(1)}(v), v)}{2} dv + (u_{i+1} - \eta_d)^2 \int_{v_j}^{v_{j+1}} \frac{R'_u(\mu_i^{(2)}(v), v)}{2} dv \\
&= (u_{i+1} - u_i)(v_{j+1} - v_j) R(\eta_d, \xi_d) + (u_{i+1} - u_i) \left( (\xi_d - v_j)^2 \frac{R'_v(\eta_d, \omega_{ij}^{(1)})}{2} + (v_{j+1} - \xi_d)^2 \frac{R'_v(\eta_d, \omega_{ij}^{(2)})}{2} \right) \\
&\quad + (\eta_d - u_i)^2 \int_{v_j}^{v_{j+1}} \frac{R'_u(\mu_i^{(1)}(v), v)}{2} dv + (u_{i+1} - \eta_d)^2 \int_{v_j}^{v_{j+1}} \frac{R'_u(\mu_i^{(2)}(v), v)}{2} dv,
\end{aligned}$$

where  $\mu_i^{(1)}(v), \mu_i^{(2)}(v) \in (u_i, u_{i+1})$  and  $\omega_{ij}^{(1)}, \omega_{ij}^{(2)} \in (v_j, v_{j+1})$ . By the mean value theorem of integral, there exist  $(\bar{\mu}_{ij}^{(1)}, \bar{\omega}_{ij}^{(1)}), (\bar{\mu}_{ij}^{(2)}, \bar{\omega}_{ij}^{(2)}) \in [u_i, u_{i+1}] \times [v_j, v_{j+1}]$  such that

$$\begin{aligned}
\int_{v_j}^{v_{j+1}} \frac{R'_u(\mu_i^{(1)}(v), v)}{2} dv &= (v_{j+1} - v_j) \frac{R'_u(\bar{\mu}_{ij}^{(1)}, \bar{\omega}_{ij}^{(1)})}{2}, \text{ and} \\
\int_{v_j}^{v_{j+1}} \frac{R'_u(\mu_i^{(2)}(v), v)}{2} dv &= (v_{j+1} - v_j) \frac{R'_u(\bar{\mu}_{ij}^{(2)}, \bar{\omega}_{ij}^{(2)})}{2},
\end{aligned}$$

Therefore,

$$\begin{aligned}
\int_{v_j}^{v_{j+1}} \int_{u_i}^{u_{i+1}} R(u, v) du dv &= (u_{i+1} - u_i)(v_{j+1} - v_j) R(\eta_d, \xi_d) \\
&\quad + (u_{i+1} - u_i) \left( (\xi_d - v_j)^2 \frac{R'_v(\eta_d, \omega_{ij}^{(1)})}{2} + (v_{j+1} - \xi_d)^2 \frac{R'_v(\eta_d, \omega_{ij}^{(2)})}{2} \right) \\
&\quad + (v_{j+1} - v_j) \left( (\eta_d - u_i)^2 \frac{R'_u(\bar{\mu}_{ij}^{(1)}, \bar{\omega}_{ij}^{(1)})}{2} + (u_{i+1} - \eta_d)^2 \frac{R'_u(\bar{\mu}_{ij}^{(2)}, \bar{\omega}_{ij}^{(2)})}{2} \right).
\end{aligned}$$

Moreover, we denote  $\Xi = [u_0, u_{n_u}] \times [v_0, v_{n_v}]$ . It is easy to show that

$$\min_{(u,v) \in \Xi} |R'_v(u, v)| \leq \sum_{j=0}^{n_v-1} \sum_{i=0}^{n_u-1} \frac{(u_{i+1} - u_i)(v_{j+1} - v_j)}{(u_{n_u} - u_0)(v_{n_v} - v_0)} \frac{|R'_v(\eta_d, \omega_{ij}^{(1)})| + |R'_v(\eta_d, \omega_{ij}^{(2)})|}{2} \leq \max_{(u,v) \in \Xi} |R'_v(u, v)|,$$

$$\min_{(u,v) \in \Xi} |R'_u(u, v)| \leq \sum_{j=0}^{n_v-1} \sum_{i=0}^{n_u-1} \frac{(u_{i+1} - u_i)(v_{j+1} - v_j)}{(u_{n_u} - u_0)(v_{n_v} - v_0)} \frac{|R'_u(\bar{\mu}_{ij}^{(1)}, \bar{\omega}_{ij}^{(1)})| + |R'_u(\bar{\mu}_{ij}^{(2)}, \bar{\omega}_{ij}^{(2)})|}{2} \leq \max_{(u,v) \in \Xi} |R'_u(u, v)|.$$

And then, based on the intermediate value theorem, there exist  $(\eta^{(1)}, \xi^{(1)}) \in \Xi$  and  $(\eta^{(2)}, \xi^{(2)}) \in \Xi$ , such that

$$\begin{aligned} |R'_v(\eta^{(1)}, \xi^{(1)})| &= \sum_{j=0}^{n_v-1} \sum_{i=0}^{n_u-1} \frac{(u_{i+1} - u_i)(v_{j+1} - v_j)}{(u_{n_u} - u_0)(v_{n_v} - v_0)} \frac{|R'_v(\eta_d, \omega_{ij}^{(1)})| + |R'_v(\eta_d, \omega_{ij}^{(2)})|}{2}, \\ |R'_u(\eta^{(2)}, \xi^{(2)})| &= \sum_{j=0}^{n_v-1} \sum_{i=0}^{n_u-1} \frac{(u_{i+1} - u_i)(v_{j+1} - v_j)}{(u_{n_u} - u_0)(v_{n_v} - v_0)} \frac{|R'_u(\bar{\mu}_{ij}^{(1)}, \bar{\omega}_{ij}^{(1)})| + |R'_u(\bar{\mu}_{ij}^{(2)}, \bar{\omega}_{ij}^{(2)})|}{2}. \end{aligned}$$

As a result,

$$\begin{aligned} \|\mathcal{DT}(u, v) - \mathcal{DT}_r(u, v)\|_{L^2}^2 &= \int_{v_0}^{v_{n_v}} \int_{u_0}^{u_{n_u}} R(u, v) du dv = \sum_{j=0}^{n_v-1} \sum_{i=0}^{n_u-1} \int_{v_j}^{v_{j+1}} \int_{u_i}^{u_{i+1}} R(u, v) du dv \\ &= \sum_{j=0}^{n_v-1} \sum_{i=0}^{n_u-1} (u_{i+1} - u_i)(v_{j+1} - v_j) R(\eta_d, \xi_d) \\ &\quad + \sum_{j=0}^{n_v-1} \sum_{i=0}^{n_u-1} (u_{i+1} - u_i) \left( (\xi_d - v_j)^2 \frac{R'_v(\eta_d, \omega_{ij}^{(1)})}{2} + (v_{j+1} - \xi_d)^2 \frac{R'_v(\eta_d, \omega_{ij}^{(2)})}{2} \right) \\ &\quad + \sum_{j=0}^{n_v-1} \sum_{i=0}^{n_u-1} (v_{j+1} - v_j) \left( (\eta_d - u_i)^2 \frac{R'_u(\bar{\mu}_{ij}^{(1)}, \bar{\omega}_{ij}^{(1)})}{2} + (u_{i+1} - \eta_d)^2 \frac{R'_u(\bar{\mu}_{ij}^{(2)}, \bar{\omega}_{ij}^{(2)})}{2} \right) \\ &\leq h^2 \sum_{d=1}^D R(\eta_d, \xi_d) + h \sum_{j=0}^{n_v-1} \sum_{i=0}^{n_u-1} (u_{i+1} - u_i)(v_{j+1} - v_j) \frac{|R'_v(\eta_d, \omega_{ij}^{(1)})| + |R'_v(\eta_d, \omega_{ij}^{(2)})|}{2} \\ &\quad + h \sum_{j=0}^{n_v-1} \sum_{i=0}^{n_u-1} (u_{i+1} - u_i)(v_{j+1} - v_j) \frac{|R'_u(\bar{\mu}_{ij}^{(1)}, \bar{\omega}_{ij}^{(1)})| + |R'_u(\bar{\mu}_{ij}^{(2)}, \bar{\omega}_{ij}^{(2)})|}{2} \\ &= h^2 e_h + h(u_{n_u} - u_0)(v_{n_v} - v_0) \left( |R'_v(\eta^{(1)}, \xi^{(1)})| + |R'_u(\eta^{(2)}, \xi^{(2)})| \right) \\ &< h^2 \bar{M} + h(u_{n_u} - u_0)(v_{n_v} - v_0) \left( |R'_v(\eta^{(1)}, \xi^{(1)})| + |R'_u(\eta^{(2)}, \xi^{(2)})| \right). \end{aligned}$$

On the other hand, since the degrees of  $u$  and  $v$  in  $T_r(u, v)$  are both larger than the maximum orders of the partial derivatives to  $u$  and  $v$  appearing in  $\mathcal{D}$  (refer to (1)), respectively, similar as the one-dimensional case,  $R'_u(u, v)$  and  $R'_v(u, v)$  are both continuous, and then bounded on  $\Omega \cup \partial\Omega$ , i.e.,

$$|R'_v(\eta^{(1)}, \xi^{(1)})| \leq \hat{M}, \text{ and, } |R'_u(\eta^{(2)}, \xi^{(2)})| \leq \hat{M},$$

where  $\hat{M}$  is a positive constant.

In conclusion,

$$\|\mathcal{DT}(u, v) - \mathcal{DT}_r(u, v)\|_{L^2}^2 \leq h^2 \bar{M} + 2h(u_{n_u} - u_0)(v_{n_v} - v_0) \hat{M},$$



and Theorem 1 is proved.  $\square$

### 3.2. Convergence

Based on the consistency property of IGA-L, we can show that IGA-L is convergent if the differential operator is stable or strongly monotonic, similarly as in Ref. [32].

Let  $V$  and  $W$  be two Hilbert spaces and  $\|\cdot\|_V$  and  $\|\cdot\|_W$  be two norms defined on  $V$  and  $W$ , respectively. Suppose  $\|\cdot\|_V$  and  $\|\cdot\|_W$  are equivalent to the  $L^2$  norm, i.e., there exist positive constants  $\alpha_V$ ,  $\beta_V$ ,  $\alpha_W$ , and  $\beta_W$  such that,

$$\alpha_V \|\cdot\|_{L^2} \leq \|\cdot\|_V \leq \beta_V \|\cdot\|_{L^2}, \quad (9)$$

$$\alpha_W \|\cdot\|_{L^2} \leq \|\cdot\|_W \leq \beta_W \|\cdot\|_{L^2}. \quad (10)$$

We first give the definitions for stable and strongly monotonic operators.

**Definition 1 (Stability estimate and stable operator [37]).** Let  $V, W$  be Hilbert spaces and  $\mathcal{D} : V \rightarrow W$  be a differential operator. If there exists a constant  $C_S > 0$  such that

$$\|\mathcal{D}v\|_W \geq C_S \|v\|_V, \quad \text{for all } v \in D(\mathcal{D}), \quad (11)$$

where  $D(\mathcal{D})$  represents the domain of  $\mathcal{D}$ , differential operator  $\mathcal{D}$  is called the stable operator, and the inequality (11) is called the stability estimate.

**Definition 2 (Strongly monotonic operator [37]).** Let  $V$  be a Hilbert space and  $\mathcal{D} \in \mathcal{L}(V, V')$ . Operator  $\mathcal{D}$  is said to be a strongly monotonic operator, if there exists a constant  $C_D > 0$ , such that

$$\langle \mathcal{D}v, v \rangle \geq C_D \|v\|_V^2, \quad \text{for all } v \in V. \quad (12)$$

For every  $v \in V$ , element  $\mathcal{D}v \in V'$  is of a linear form. The symbol  $\langle \mathcal{D}v, v \rangle$ , which denotes the application of  $\mathcal{D}v$  to  $v \in V$ , is called a duality pairing.

Clearly, if a differential operator  $\mathcal{D}$  is strongly monotonic, it is stable.

**Lemma 2.** Let  $V$  be a Hilbert space and  $\mathcal{D} \in \mathcal{L}(V, V')$  be a continuous strongly monotonic linear operator. Then, there exists a constant  $C_D > 0$  such that  $\mathcal{D}$  satisfies the stability estimate (11) [37].

The proof can be found in Ref. [32].

Therefore, we have the convergence property of IGA-L as follows.

**Theorem 2.** Suppose NURBS function  $T_r$ , defined on knot grid  $\mathcal{T}^h$ , is the numerical solution to the boundary value problem (1), generated by IGA-L, and the conditions presented in Theorem 1 are satisfied in one, two, and three dimensions, respectively. If differential operator  $\mathcal{D} : V \rightarrow W$  in (1) is a stable operator,  $T_r$  will converge to analytic solution  $T$  when the knot grid size  $h \rightarrow 0$ .

**Proof:** Differential operator  $\mathcal{D}$  in (1) is a stable operator, so there exists a constant  $C_S > 0$ , such that

$$\|\mathcal{D}(T - T_r)\|_W \geq C_S \|T - T_r\|_V.$$

And it is equivalent to

$$\|T - T_r\|_V \leq \frac{1}{C_S} \|\mathcal{D}T - \mathcal{D}T_r\|_W.$$

Due to the equivalence of  $\|\cdot\|_{L^2}$  and  $\|\cdot\|_W$  (9), we have,

$$\|T - T_r\|_V \leq \frac{1}{C_S} \|\mathcal{D}T - \mathcal{D}T_r\|_W \leq \frac{\beta_W}{C_S} \|\mathcal{D}T - \mathcal{D}T_r\|_{L^2}.$$

Because of the consistency of IGA-L (Theorem 1),  $\|T - T_r\|_V$  will converge to 0 when  $h \rightarrow 0$ . And this theorem is proved.  $\square$

Moreover, Theorem 2 and Lemma 2 lead to the direct corollary.

**Corollary 1.** *Suppose NURBS function  $T_r$  defined on knot grid  $\mathcal{T}^h$  is the numerical solution to the boundary value problem (1), generated by IGA-L, and the conditions presented in Theorem 1 are satisfied in one, two, and three dimensions, respectively. Additionally, suppose norm  $\|\cdot\|_{L^2}$  bounds norm  $\|\cdot\|_V$ . If differential operator  $\mathcal{D}$  in (1) is a strongly monotonic operator, then  $T_r$  will converge to analytic solution  $T$  when  $h \rightarrow 0$ .*

It is well known that a wide class of elliptic differential operators are stable or strongly monotonic. Hence, IGA-L is convergent for equations that have these elliptic differential operators. The examples of a PDE whose differential operators are strongly monotonic can be found in Refs. [37, 32].

## 4. Comparisons with IGA-C

### 4.1. Theoretical comparisons

In the following, we compare IGA-L with IGA-C in terms of their computational efficiency at solving a scalar problem (Laplace equation [31]) and vector problem (elasticity equation [31]). We consider model discretizations in one, two, and three dimensions that are characterized by the degree of the basis functions and the numbers of control and collocation points in each parametric direction. For the sake of simplicity, we assume that the model discretizations in two and three dimensions have the same number of collocation points (and control points) in each parametric direction, and the degrees of basis functions in one, two, and three dimensions are  $p$ ,  $p \times p$ , and  $p \times p \times p$ , respectively.

First, the costs in floating point operations (flops) for the formation at one collocation point are the same for both IGA-L and IGA-C. These costs are listed in Table 1. It should be pointed out that there are some slight flaws in the formulae employed to calculate the costs in flops in Ref. [31]. Hence, the formulae presented in Table 1 are a bit different from those in [31].

Second, we compared the costs (in flops) of solving the linear system of equations in IGA-C and IGA-L, and present the comparison in Table 2. In this table, the first column is the dimension of the problem solved, the second column is the number of control points (equal to the number of collocation points in IGA-C), and the third column is the cost in flops to solve the linear system of equations using Gaussian elimination in IGA-C [36]. In addition, the fourth column is the number of control points, and the fifth column is the number of collocation points in IGA-L. Finally, the last column is the cost in flops to solve normal equation (5) using Cholesky decomposition in IGA-L [36]. We can see that the cost to solve Eq. (5) in IGA-L linearly increases with the number of collocation points ( $m$ ,  $m \times m$ ,  $m \times m \times m$ ).

Table 1: Cost in flops for formation at one collocation point in IGA-C and IGA-L.

Dimension	A scalar problem (Laplace)	A vector problem (elasticity)
Solve for 1st derivatives		
1		$(p + 1)$
2		$5(p + 1)^2 + 4$
3		$12(p + 1)^3 + 16$ [20]
Compute right hand side vectors and solve for 2nd derivatives		
1		$3(p + 1)$
2		$24(p + 1)^2 + 16$ [20]
3		$87(p + 1)^3 + 140$
Total number of flops for basis function		
1		$35(p + 1) + 1$ [2]
2		$124(p + 1)^2 + 33$ [37]
3		$302(p + 1)^3 + 219$ [223]
		Evaluate Navier's eqs. on global level
1		$(p + 1)$
2		$12(p + 1)^2$
3		$21(p + 1)^3$
Total number of flops to evaluate the local stiffness matrix		
1	$35(p + 1) + 1$ [2]	$36(p + 1) + 1$ [2]
2	$125(p + 1)^2 + 33$ [37]	$134[136](p + 1)^2 + 33$ [37]
3	$304(p + 1)^3 + 219$ [223]	$323(p + 1)^3 + 219$ [223]

Note: The formulae in this Table is a bit different from those in Ref. [31], and the numbers in the square brackets are provided by Ref. [31].

Table 2: Cost comparison of IGA-C with IGA-L.

Dimension	IGA-C		IGA-L		
	Num. <sup>1</sup>	Cost in flops.	Num. <sup>1</sup>	Num. <sup>2</sup>	Cost in flops.
$d = 1$	$n$	$2n^3/3$	$n$	$m$	$mn^2 + n^3/3$
$d = 2$	$n \times n$	$2n^6/3$	$n \times n$	$m \times m$	$m^2n^4 + n^6/3$
$d = 3$	$n \times n \times n$	$2n^9/3$	$n \times n \times n$	$m \times m \times m$	$m^3n^6 + n^9/3$

<sup>1</sup> Number of control points.

<sup>2</sup> Number of collocation points.

#### 4.2. Numerical comparisons

In this section, we compare IGA-L with IGA-C using some numerical examples. To measure the approximation accuracy, we define two error formulae, i.e., the *relative error for the solution*  $T_r$ ,

$$e_T = \sqrt{\frac{\int_{\Omega} (T - T_r)'(T - T_r) d\Omega}{\int_{\Omega} T' T d\Omega}}, \quad (13)$$

and the *relative error for the differential operator*  $\mathcal{DT}_r$ ,

$$e_{\mathcal{DT}} = \sqrt{\frac{\int_{\Omega} (\mathcal{DT} - \mathcal{DT}_r)' (\mathcal{DT} - \mathcal{DT}_r) d\Omega}{\int_{\Omega} (\mathcal{DT})' (\mathcal{DT}) d\Omega}}. \quad (14)$$

Additionally, to illustrate the error distribution of the numerical solution, the following *absolute errors*  $e_a$  are employed, i.e.,

$$\begin{aligned} e_a(u) &= |T(u) - T_r(u)|, \text{ for one-dimensional case,} \\ e_a(u, v) &= |T(u, v) - T_r(u, v)|, \text{ for two-dimensional case,} \\ e_a(u, v, w) &= |T(u, v, w) - T_r(u, v, w)|, \text{ for three-dimensional case.} \end{aligned} \quad (15)$$

In the following, five numerical examples are presented. These examples are implemented with MATLAB and run on a PC with a 2.66-GHz Intel Core2 Quad CPU Q9400 and 3 GB memory. Examples I–III are three source problems in one, two, and three dimensions, respectively. Example IV is a linear elasticity problem, and Example V demonstrates the stability of the IGA-L method with respect to that of the IGA-C method. The problems in Examples I–IV are solved by both the IGA-C and IGA-L methods. With the IGA-C method, the control points are increased gradually, and the collocation points are the Greville abscissae [2] of the knot vectors, also called the *Greville collocation points*. With the IGA-L method, we employed two strategies. In the first strategy, named **IGA-L fixed**, the control points are fixed and the collocation points are increased gradually at the same rate as the collocation points used in the corresponding IGA-C. In the second strategy, named **IGA-L variable**, the control points are variable and are increased gradually at the same rate as those of IGA-C. In each computation round of the IGA-L variable strategy, supposing the numbers of the control points are  $n$ ,  $n \times n$ , and  $n \times n \times n$  in the one-, two-, and three-dimensional cases, respectively, the numbers of the corresponding collocation points are taken as  $n + 2$ ,  $(n + 2) \times (n + 2)$ , and  $(n + 2) \times (n + 2) \times (n + 2)$ , respectively. The collocation points for IGA-L are also taken as the Greville abscissae of the knot vectors. To produce the Greville collocation points, we repeatedly inserted the knot at the midpoint of the longest knot interval until there were enough knots in the knot vector.

The plots for the numerical results of Examples I–III are illustrated in Figs. 2, 4, and 6, respectively. In each of these figures, sub-figures (a) and (d) are plots of the logarithm (base 10) of the relative error for the solution vs. the number of collocation points, (b) and (e) are plots of the logarithm (base 10) of the relative error for the differential operator vs. number of collocation points, and (c) and (f) are plots of the running time vs. number of collocation points. Moreover, in these figures, the blue, red, and light green lines indicate the data for the IGA-C, IGA-L fixed, and IGA-L variable strategies, respectively.

**Example I:** one-dimensional source problem with Dirichlet boundary condition:

$$\begin{cases} -T'' + T = (1 + 4\pi^2) \sin(2\pi x), & x \in \Omega = [0, 1], \\ T(0) = 0, T(1) = 0. \end{cases} \quad (16)$$

This problem's analytical solution is  $T = \sin(2\pi x)$ . The physical domain  $\Omega = [0, 1]$  of the boundary problem (16) is represented by a cubic B-spline curve. For the mathematical representation of the B-spline curve, please refer to Appendix A1. The B-spline curves with additional control points utilized in this example were generated by uniformly inserting knots to the knot vector of the cubic B-spline curve presented in Appendix A1.

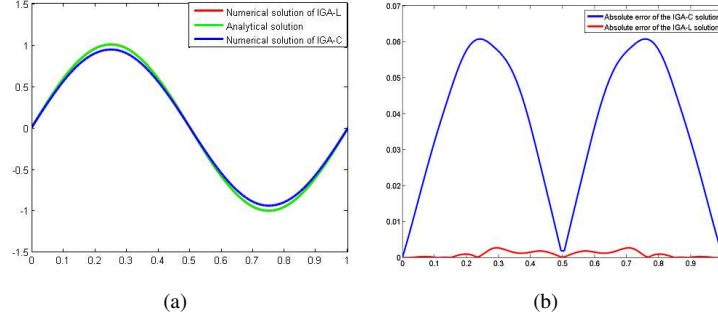


Figure 1: Comparison of the analytical solution, IGA-L solution, and IGA-C solution of Eq. (16). (a) Analytical solution, IGA-L, and IGA-C solution. Note that the analytical solution almost overlaps the IGA-L solution. (b) Absolute error distribution curves of the IGA-L and IGA-C solutions.

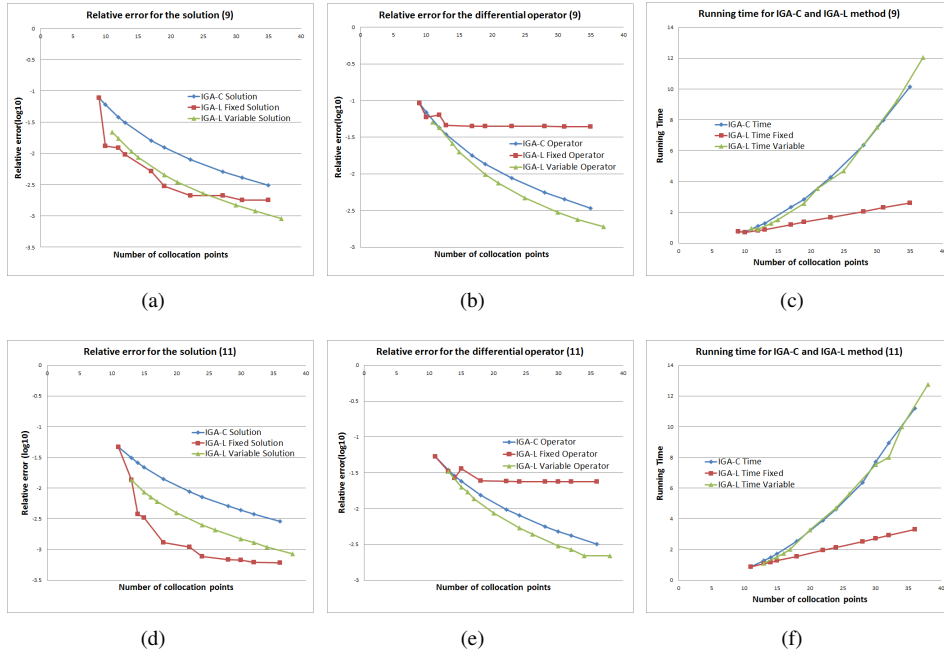


Figure 2: Numerical results for the 1D problem (16). In (a), (b), and (c), 9 control points are fixed in the IGA-L fixed strategy; in (d), (e), and (f), 11 control points are fixed in the IGA-L fixed strategy. (a) and (d) Relative error of the solution vs. number of collocation points. (b) and (e) Relative error of the differential operator vs. number of collocation points. (c) and (f) Running time vs. number of collocation points.

The analytical, IGA-C, and IGA-L solutions for the one-dimensional source problem (16), are illustrated in Fig. 1(a), where the IGA-C solution was generated with 10 control points, and the IGA-L solution was generated with 10 control points and 16 Greville collocation points. The relative errors for the IGA-C and IGA-L solutions are 0.0598 and 0.0018, respectively. The relative error for the IGA-L solution is one order of magnitude less than that of the IGA-C

solution. In addition, Fig. 1(b) demonstrates the absolute error distribution curves of the IGA-C and IGA-L solutions, respectively. The maximum absolute errors of the IGA-C and IGA-L solutions are 0.0607 and 0.0028, respectively.

Moreover, plots for the numerical results of the one-dimensional problem (16) are displayed in Fig. 2. In Figs. 2(a), 2(b), and 2(c), 9 control points are fixed in the IGA-L fixed strategy; in Figs. 2(d), 2(e), and 2(f), 11 control points are fixed in the IGA-L fixed strategy.

As shown in Figs. 2(a) and 2(d) (the red lines), although the control points are fixed in the IGA-L fixed strategy, the relative error of the solution decreases as the number of collocation points increases. For the first several increases, the relative error reduces significantly. Afterwards, it tends to a steady value. On the other hand, as illustrated by the light green lines, the relative error of the solution produced by the IGA-L variable strategy is always less than that of the IGA-C solution (blue lines) with the same control points, and they follow nearly the same declining trend.

Figs. 2(b) and 2(e) illustrate the data for the differential operator. We also note that while the relative errors for the differential operators generated by the IGA-C and IGA-L variable strategies decrease gradually with the increase of control points, the relative error for the operator generated by the IGA-L fixed strategy (which has fixed control points) tends to a stable value (indicated by the red line).

Finally, the data for the running time is illustrated in Figs. 2(c) and 2(f). It can be seen that the IGA-C and IGA-L variable strategies have nearly the same plots (blue and light green lines), and the increase in the running time of the IGA-L fixed strategy (red line) is less than those of the IGA-C and IGA-L variable strategies.

As a result, with the IGA-L fixed strategy, where the control points are fixed and the collocation points increase, the relative error decreases and then tends to a stable value. The relative error of the IGA-L variable strategy is always less than that of IGA-C, though their running times are nearly the same.

**Example II:** source problem in the two-dimensional domain  $\Omega$ ,

$$\begin{cases} -\Delta T + T = f, & (x, y) \in \Omega \\ T|_{\partial\Omega} = 0, \end{cases} \quad (17)$$

where  $\Omega$  is a quarter of an annulus, which can be exactly represented by a cubic NURBS patch with  $4 \times 4$  control points, as presented in Appendix A2, and where

$$\begin{aligned} f = & (3x^4 - 67x^2 - 67y^2 + 3y^4 + 6x^2y^2 + 116) \sin(x) \sin(y) \\ & + (68x - 8x^3 - 8xy^2) \cos(x) \sin(y) \\ & + (68y - 8y^3 - 8yx^2) \cos(y) \sin(x). \end{aligned}$$

The analytical solution of the source problem (17) is

$$T = (x^2 + y^2 - 1)(x^2 + y^2 - 16) \sin(x) \sin(y).$$

We note that the cubic NURBS patches with additional control points that represent the two-dimensional domain  $\Omega$  in this example were generated by uniformly inserting knots on the cubic NURBS patch with  $4 \times 4$  control points, as described in Appendix A2.

Fig. 3 presents the analytical solution and numerical solutions of the two-dimensional source problem (17), as generated by the IGA-C and IGA-L methods. To produce the numerical solutions, we uniformly inserted 11 knots along the  $u$ - and  $v$ - directions, respectively, to the cubic

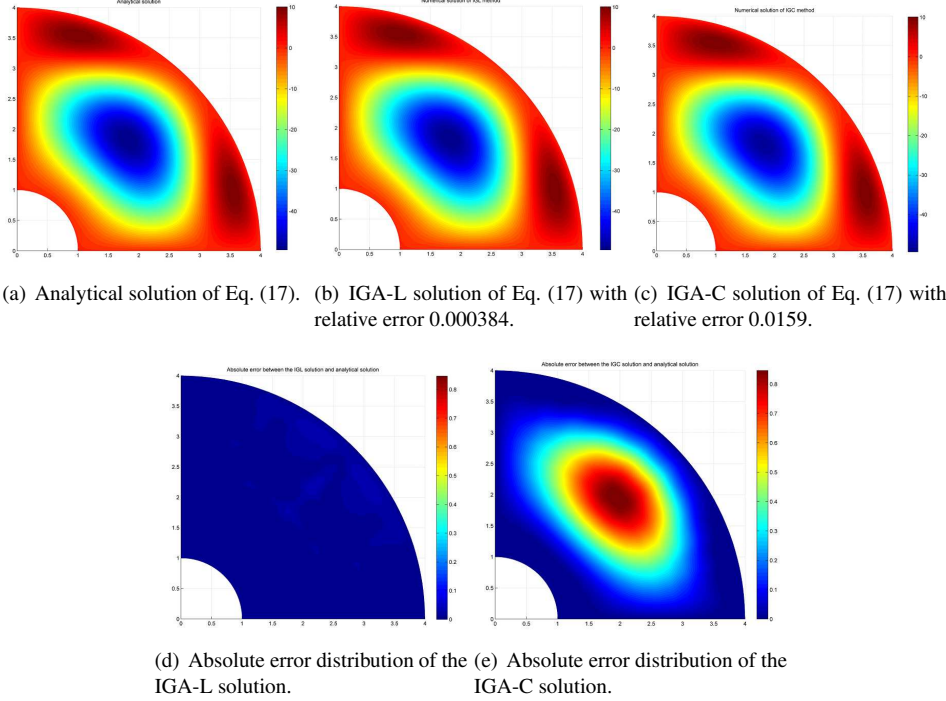


Figure 3: Comparison of the analytical solution, IGA-L solution, and IGA-C solution of Eq. (17). The relative error of the IGA-L solution is nearly two orders of magnitude less than the IGA-C solution.

NURBS patch presented in Appendix A2, resulting in a cubic NURBS patch with  $15 \times 15$  control points. With  $20 \times 20$  Greville collocation points, the IGA-L method was employed to solve Eq. (17). The relative error of the IGA-L solution (see Fig. 3(b)) is 0.000384, and Fig. 3(d) illustrates the absolute error distribution of the IGA-L solution. In addition, the source problem (17) was also solved by the IGA-C method using the same NURBS patch of  $15 \times 15$  control points (see Fig. 3(c)) with Greville collocation points. The relative error of the IGA-C solution is 0.0159, and its absolute error distribution is illustrated in Fig. 3(e). In this example, the relative error of the IGA-L solution is two orders of magnitude less than that of the IGA-C solution.

Moreover, the plots of the numerical results for Example II are illustrated in Fig. 4. In Figs. 4(a), 4(b), and 4(c), the number of control points in the IGA-L fixed strategy is fixed at  $7 \times 7$ ; in Figs. 4(d), 4(e), and 4(f), the number of the control points in the IGA-L fixed strategy is fixed at  $9 \times 9$ . Based on the numerical results illustrated in Fig. 4, we can draw similar conclusions as in Example I. For the IGA-L fixed strategy, while the number of collocation points increases, the relative errors for both the solution and differential operator first decrease, and then tend to a stable value. For the IGA-L variable strategy, the relative error for the solution is always less than that of the corresponding IGA-C solution. Furthermore, the plots for the running time vs. number of collocation points of the three methods are nearly the same.

**Example III:** source problem defined on the three-dimensional cubic domain  $\Omega = [0, 1] \times$

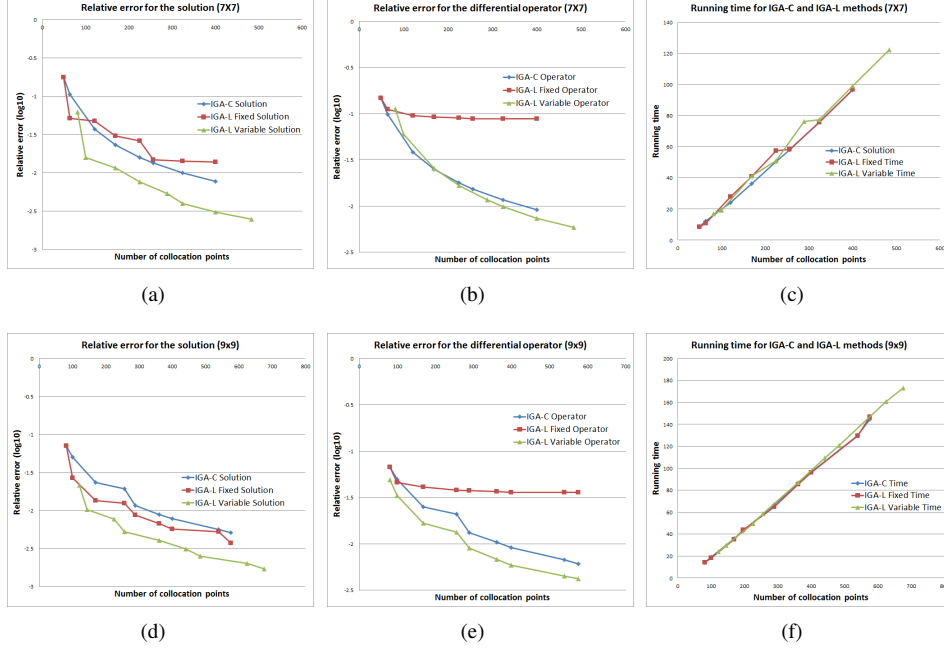


Figure 4: Numerical results for the 2D problem (17). In (a), (b), and (c),  $7 \times 7$  control points were fixed in the IGA-L fixed strategy; in (d), (e), and (f),  $9 \times 9$  control points were fixed in the IGA-L fixed strategy. (a) and (d) Relative error for the solution vs. number of collocation points. (b) and (e) Relative error for the differential operator vs. number of collocation points. (c) and (f) Running time vs. number of collocation points.

$[0, 1] \times [0, 1]$ , i.e.,

$$\begin{cases} -\Delta T + T = f, & (x, y, z) \in \Omega, \\ T|_{\partial\Omega} = 0, \end{cases} \quad (18)$$

where

$$f = (1 + 12\pi^2) \sin(2\pi x) \sin(2\pi y) \sin(2\pi z).$$

Its analytical solution is,

$$T = \sin(2\pi x) \sin(2\pi y) \sin(2\pi z).$$

The three-dimensional physical domain  $\Omega$  is modeled as a cubic B-spline solid with  $4 \times 4 \times 4$  control points, as listed in Appendix A3. Furthermore, the cubic B-spline solids with additional control points employed in this example were generated by uniformly inserting knots to the B-spline solid in Appendix A3.

In Fig. 5, the analytical solution, IGA-L solution, and IGA-C solution for Eq. (18) are illustrated, where the IGA-L and IGA-C solutions were generated with  $7 \times 7 \times 7$  control points. Specifically, to get the IGA-L solution,  $10 \times 10 \times 10$  Greville collocation points were employed, and the relative error is 0.0232 (see Fig. 5(b)). On the other hand, the relative error for the IGA-C solution with Greville collocation points is 0.1456 (Fig. 5(c)). In this example, the relative error of the IGA-L solution is one order of magnitude smaller than that of the IGA-C solution. Additionally, Figs. 5(d) and 5(e) present the absolute error distribution for the IGA-L and IGA-C solutions, respectively.



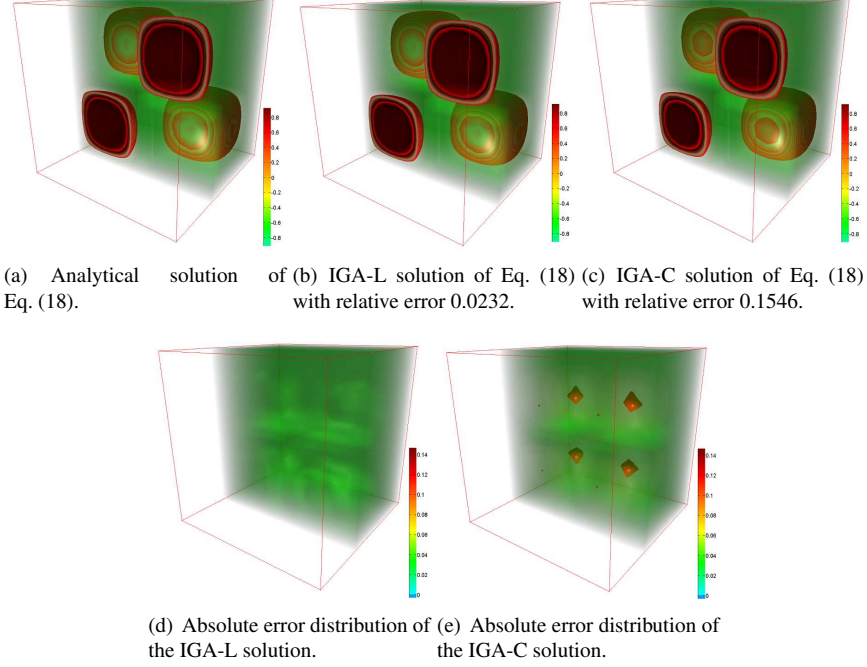


Figure 5: Comparison of the analytical solution, IGA-L solution, and IGA-C solution of Eq. (18). The relative error of the IGA-L solution is nearly one order of magnitude less than that of the IGA-C solution.

Furthermore, plots of the numerical results for Eq. (18) are illustrated in Fig. 6. From Fig. 6, we can see that the performances of the IGA-C, IGA-L fixed strategy, and IGA-L variable strategy are similar to those in the one- and two-dimensional examples.

**Example IV:** elastic problem, i.e., the simply supported beam (see Fig. 7). As illustrated in Fig. 7, the simply supported beam with a rectangular cross section has depth  $h$  and length  $2l$ . Uniformly distributed loading  $q$  was applied on the upper surface, and equilibrium was maintained by reaction force  $ql$  at both ends. Here, the body force need not be considered. The analytical solution of the simply supported beam problem is

$$\begin{aligned}\sigma_x &= \frac{6q}{h^3}(l^2 - x^2)y + q\frac{y}{h}\left(4\frac{y^2}{h^2} - \frac{3}{5}\right), \\ \sigma_y &= -\frac{q}{2}\left(1 + \frac{y}{h}\right)\left(1 - \frac{2y}{h}\right)^2, \\ \tau_{xy} &= -\frac{6q}{h^3}x\left(\frac{h^2}{4} - y^2\right).\end{aligned}$$

We calculated the simply supported beam problem using the IGA-C, IGA-L fixed, and IGA-L variable methods, with  $q = 10$ ,  $h = 2$ ,  $l = 5$  (see Fig. 7). The physical domain was represented by a cubic B-spline patch presented in Appendix A4. Fig. 8 illustrates the analytical and numerical solutions for  $\sigma_x$ ,  $\sigma_y$ , and  $\tau_{xy}$ , generated by the IGA-L method with  $11 \times 11$  control points and  $18 \times 18$  Greville collocation points. The relative errors for the IGA-L solution for  $\sigma_x$ ,  $\sigma_y$ , and  $\tau_{xy}$

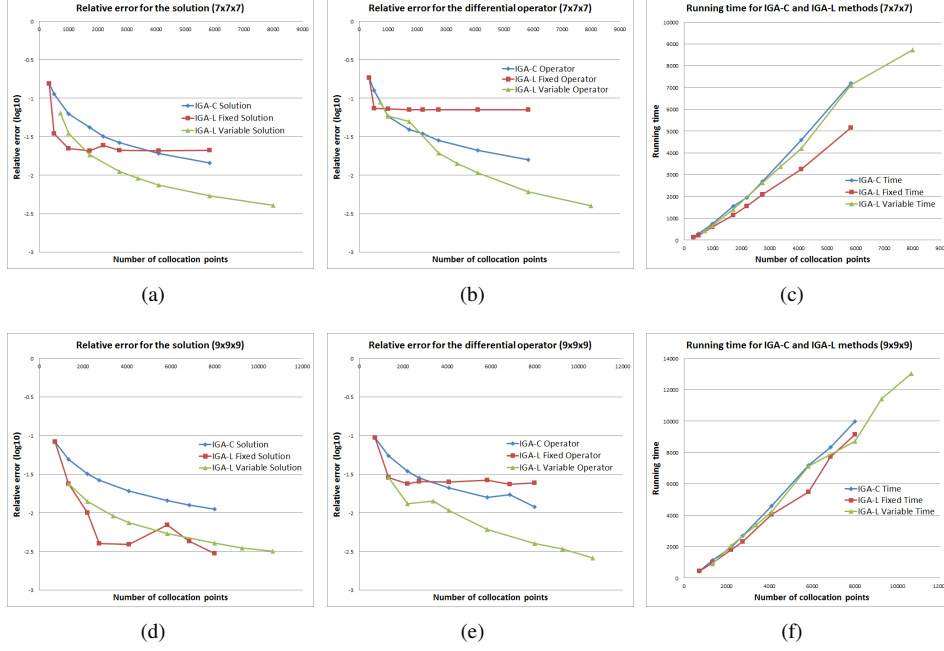


Figure 6: Numerical results for the 3D source problem (18). In (a), (b), and (c),  $7 \times 7 \times 7$  control points were fixed in the IGA-L fixed strategy; in (d), (e), and (f),  $9 \times 9 \times 9$  control points were fixed in the IGA-L fixed strategy. (a) and (d) Relative error for the solution vs. number of collocation points. (b) and (e) Relative error for the differential operator vs. number of collocation points. (c) and (f) Running time vs. number of collocation points.

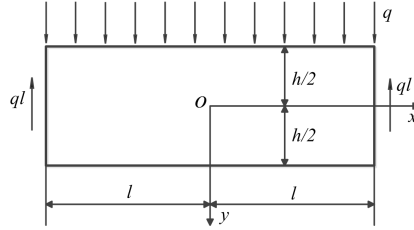


Figure 7: The simply supported beam.

are  $1.1 \times 10^{-5}$ ,  $3.29 \times 10^{-4}$ , and  $7.2 \times 10^{-5}$ , respectively. Moreover, Fig. 9 demonstrates the absolute error distribution for the IGA-C solution with  $11 \times 11$  control points and the IGA-L solution with  $11 \times 11$  control points and  $18 \times 18$  Greville collocation points. While the maximum absolute errors for the IGA-C solution  $\sigma_x$ ,  $\sigma_y$ , and  $\tau_{xy}$  are 2.3701, 0.7329, and 1.0174, respectively, the maximum absolute errors for the IGA-L solution  $\sigma_x$ ,  $\sigma_y$ , and  $\tau_{xy}$  are just 0.0040, 0.0112, and 0.0053, respectively.

Furthermore, in Fig. 10, we plot the relative error of the numerical solution vs. number of collocation points and running time vs. number of collocation points. We note that the IGA-C solutions do not show convergence as the number of control points increases, as illustrated in Figs. 10(a)– 10(c) (blue lines). However, the relative errors of the IGA-L fixed strategy decrease

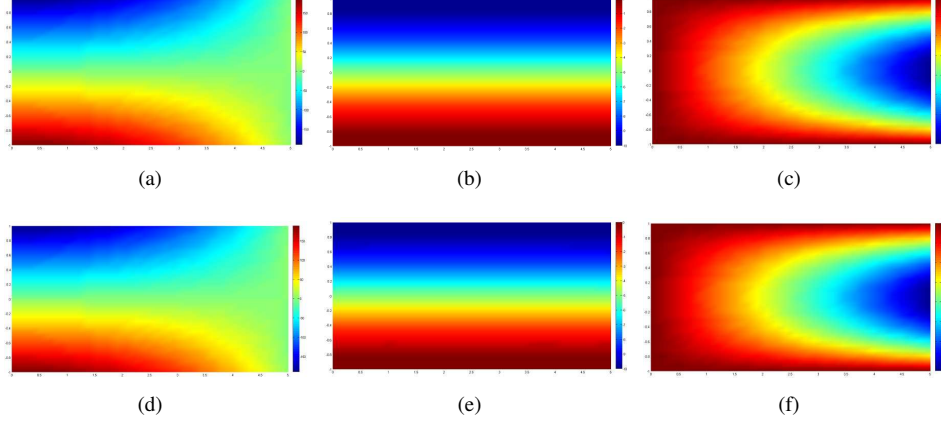


Figure 8: Analytical solution for  $\sigma_x$  (a),  $\sigma_y$  (b), and,  $\tau_{xy}$  (c) and numerical solution for  $\sigma_x$  (d),  $\sigma_y$  (e), and,  $\tau_{xy}$  (f) generated by the IGA-L method with  $11 \times 11$  control points and  $18 \times 18$  Greville collocation points.

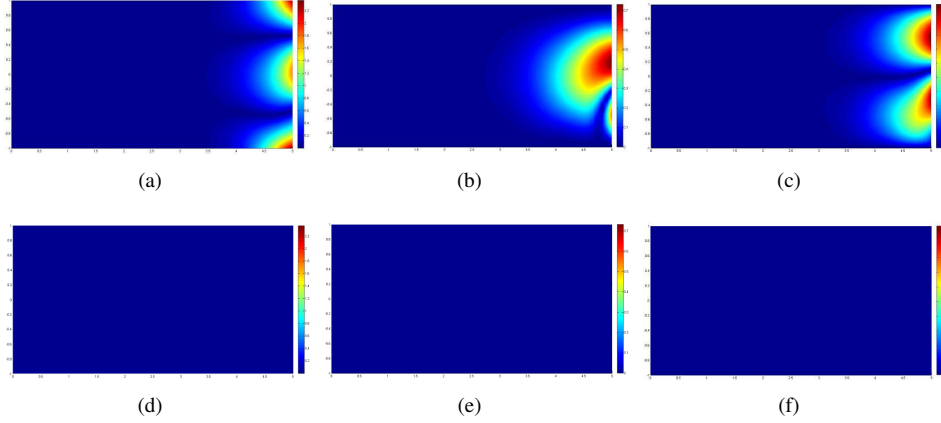


Figure 9: Absolute error (15) distribution for the IGA-C solution  $\sigma_x$  (a),  $\sigma_y$  (b), and  $\tau_{xy}$  (c) and the IGA-L solution  $\sigma_x$  (d),  $\sigma_y$  (e), and  $\tau_{xy}$  (f).

as the number of collocation points increases (red lines). Additionally, although the running time of the IGA-C method rapidly rises with more and more control points, the running time of the IGA-L fixed strategy increases at a nearly linear rate (Fig. 10(d)).

Note that in the IGA-C method the number of collocation points is fixed to be equal to the number of control points. However, in the IGA-L method, the number of collocation points is variable and larger than the number of control points. Therefore, the IGA-L method is more flexible than the IGA-C method. Consequently, in the cases where the IGA-C method is unstable, the IGA-L method can be made stable by changing the number of collocation points, as shown in the last example.

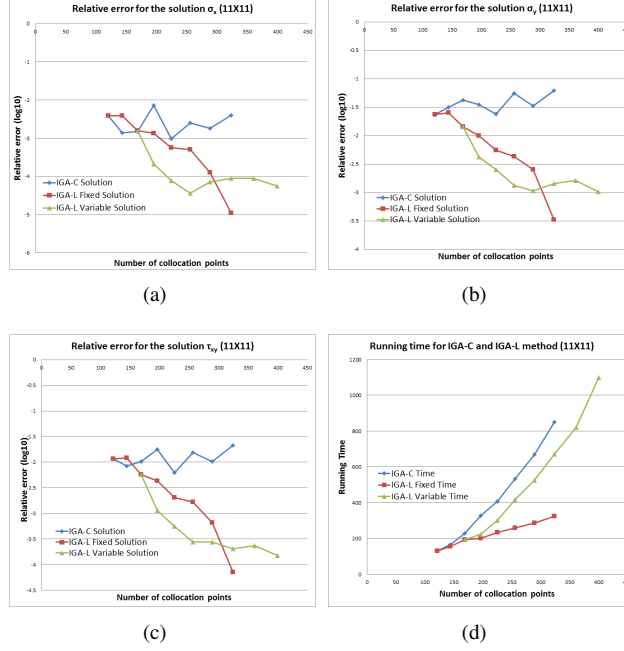


Figure 10: Numerical results for the simply supported beam (Fig. 7). In the IGA-L fixed strategy,  $11 \times 11$  control points were fixed. (a) Relative error for solution  $\sigma_x$  vs. number of collocation points. (b) Relative error for solution  $\sigma_y$  vs. number of collocation points. (c) Relative error for solution  $\tau_{xy}$  vs. number of collocation points. (d) Running time vs. number of collocation points.

**Example V:** The last example is also a one-dimensional source problem,

$$\begin{cases} -T'' + T = (1 + 4\pi^2) \sin(2\pi x), & x \in \Omega = [0, 1], \\ T(0) = 0, \\ T'(1) = 2\pi, \end{cases} \quad (19)$$

where there is a Dirichlet boundary condition at the left end, a Neumann boundary condition at the right end, and the PDE is the same as that of the boundary value problem (16). The analytical solution of Eq. (19) is  $T = \sin(2\pi x)$ . We still employ the cubic B-spline curve, presented in Appendix A1, to represent the physical domain  $\Omega = [0, 1]$ .

**Stability:** In this example, we show that the IGA-L method can be made stable by choosing appropriate number of collocation points in the cases where the IGA-C method is unstable. Consider the case where the analytic solution of Eq. (19) is approximated by a cubic B-spline function with 10 unknown control points, generated by inserting the following knots

$$0.25, 0.5, 0.6, 0.7, 0.75, 0.8,$$

into the cubic B-spline curve in Appendix A1. When we use IGA-C to solve the source problem (19) both with uniform collocation points and Greville collocation points, the numerical solutions are unstable, with relative errors (13)  $1.3999 \times 10^4$  (Fig. 11(b)) and  $2.6245 \times 10^3$  (Fig. 11(d)),

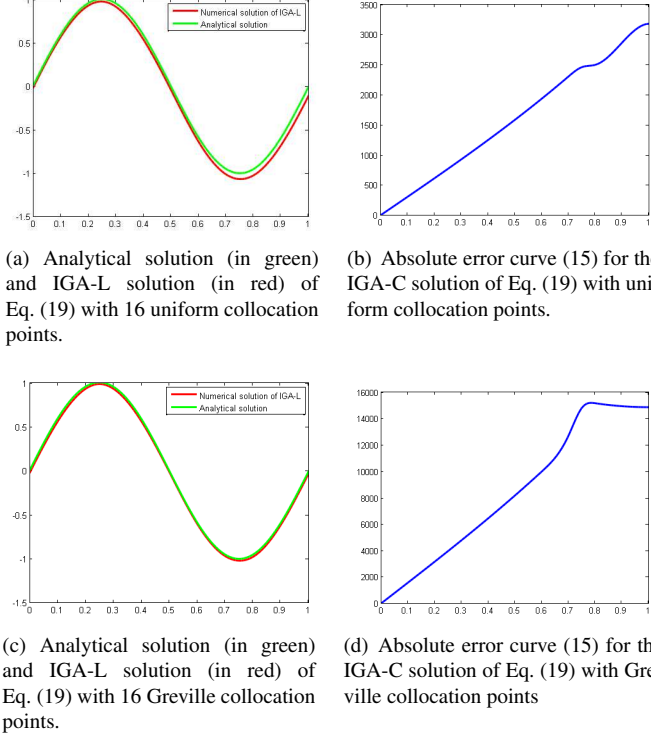


Figure 11: IGA-L can be made more stable than IGA-C by changing the number of collocation points when solving Eq. (19). (a) The IGA-L solution to Eq. (19) with 16 uniform collocation points is stable, and the relative error is 0.0805. (b) The IGA-C solution to Eq. (19) with uniform collocation points is unstable, and the relative error is  $1.3999 \times 10^4$ . (c) The IGA-L solution to Eq. (19) with 16 Greville collocation points is stable, and the relative error is 0.0343. (d) The IGA-C solution to Eq. (19) with Greville collocation points is unstable, and the relative error is  $2.6245 \times 10^3$ .

respectively. However, when IGA-L is employed to solve the source problem (19) with 16 collocation points, both with uniform collocation points and Greville collocation points, the solutions are stable, with relative errors 0.0805 (Fig. 11(a)), and 0.0343 (Fig. 11(c)), respectively.

## 5. Conclusion

IGA approximates the solution of a boundary value problem (or initial value problem) by a NURBS function. In this paper, we developed the IGA-L method to determine the unknown coefficients of the approximate NURBS function by fitting the sampling values in the least-squares sense. We proved the consistency and convergence properties of IGA-L. Moreover, the many numerical examples presented in this paper show that a small computational increase in IGA-L leads to large improvements in the accuracy, and furthermore that IGA-L is more flexible and more stable than IGA-C.

## Acknowledgement

This work is supported by the National Natural Science Foundation of China (Grant Nos. 61379072, 61202201, 60970150). Dr. Qianqian Hu is also supported by the Open Project Program (No. A1305) of the State Key Lab of CAD&CG, Zhejiang University.

## Appendix

In the Appendix, we list the control points, knot vector, and weights of the NURBS representation of the physical domains in the numerical examples.

### A1: NURBS representation of the physical domain in Examples I and IV

The physical domain in Examples I and IV is represented by a cubic B-spline curve. Its control points are listed in the following Table 3.

Table 3: Control points of the cubic B-spline curve in Examples I and IV

$B_1$	$B_2$	$B_3$	$B_4$
0	$\frac{1}{3}$	$\frac{2}{3}$	1

The knot vector is

$$0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1.$$

### A2: NURBS representation of the physical domain in Example II

The physical domain in Example II is represented by a cubic NURBS patch. Its control points are listed in Table 4.

Table 4: Control points of the quarter of annulus

$i$	$B_{i,1}$	$B_{i,2}$	$B_{i,3}$	$B_{i,4}$
1	(1,0)	(2,0)	(3,0)	(4,0)
2	$(1, 2-\sqrt{2})$	$(2, 4-2\sqrt{2})$	$(3, 6-3\sqrt{2})$	$(4, 8-4\sqrt{2})$
3	$(2-\sqrt{2}, 1)$	$(4-2\sqrt{2}, 2)$	$(6-3\sqrt{2}, 3)$	$(8-4\sqrt{2}, 4)$
4	(0,1)	(0,2)	(0,3)	(0,4)

Table 5 presents its weights.

Table 5: Weights for the quarter of annulus

$i$	$\omega_{i,1}$	$\omega_{i,2}$	$\omega_{i,3}$	$\omega_{i,4}$
1	1	1	1	1
2	$\frac{1+\sqrt{2}}{3}$	$\frac{1+\sqrt{2}}{3}$	$\frac{1+\sqrt{2}}{3}$	$\frac{1+\sqrt{2}}{3}$
3	$\frac{1+\sqrt{2}}{3}$	$\frac{1+\sqrt{2}}{3}$	$\frac{1+\sqrt{2}}{3}$	$\frac{1+\sqrt{2}}{3}$
4	1	1	1	1

The knot vectors along  $u$ - and  $v$ -direction are, respectively,

$$0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1,$$

$$0 \ 0 \ 0 \ 0 \ 1 \ 1 \ 1 \ 1.$$

*A3: NURBS representation of the physical domain in Example III*

The physical domain in Example III is represented by a cubic B-spline solid. Its control points are listed in Table 6.

Table 6: Control points of the cubic B-spline solid in Example III					
$i$	$j$	$B_{i,j,1}$	$B_{i,j,2}$	$B_{i,j,3}$	$B_{i,j,4}$
1	1	(0,0,0)	(0,0,1/3)	(0,0,2/3)	(0,0,1)
1	2	(0,1/3,0)	(0,1/3,1/3)	(0,1/3,2/3)	(0,1/3,1)
1	3	(0,2/3,0)	(0,2/3,1/3)	(0,2/3,2/3)	(0,2/3,1)
1	4	(0,1,0)	(0,1,1/3)	(0,1,2/3)	(0,1,1)
2	1	(1/3,0,0)	(1/3,0,1/3)	(1/3,0,2/3)	(1/3,0,1)
2	2	(1/3,1/3,0)	(1/3,1/3,1/3)	(1/3,1/3,2/3)	(1/3,1/3,1)
2	3	(1/3,2/3,0)	(1/3,2/3,1/3)	(1/3,2/3,2/3)	(1/3,2/3,1)
2	4	(1/3,1,0)	(1/3,1,1/3)	(1/3,1,2/3)	(1/3,1,1)
3	1	(2/3,0,0)	(2/3,0,1/3)	(2/3,0,2/3)	(2/3,0,1)
3	2	(2/3,1/3,0)	(2/3,1/3,1/3)	(2/3,1/3,2/3)	(2/3,1/3,1)
3	3	(2/3,2/3,0)	(2/3,2/3,1/3)	(2/3,2/3,2/3)	(2/3,2/3,1)
3	4	(2/3,1,0)	(2/3,1,1/3)	(2/3,1,2/3)	(2/3,1,1)
4	1	(1,0,0)	(1,0,1/3)	(1,0,2/3)	(1,0,1)
4	2	(1,1/3,0)	(1,1/3,1/3)	(1,1/3,2/3)	(1,1/3,1)
4	3	(1,2/3,0)	(1,2/3,1/3)	(1,2/3,2/3)	(1,2/3,1)
4	4	(1,1,0)	(1,1,1/3)	(1,1,2/3)	(1,1,1)

The knot vectors along  $u$ -,  $v$ -, and  $w$ -directions are, respectively,

$$\begin{aligned} &0\ 0\ 0\ 0\ 1\ 1\ 1\ 1, \\ &0\ 0\ 0\ 0\ 1\ 1\ 1\ 1, \\ &0\ 0\ 0\ 0\ 1\ 1\ 1\ 1. \end{aligned}$$

*A4: NURBS representation of the physical domain in Example IV*

The physical domain in Example IV is represented by a cubic B-spline patch. Its control points are listed in Table 7.

Table 7: Control points of the quarter of annulus				
$i$	$B_{i,1}$	$B_{i,2}$	$B_{i,3}$	$B_{i,4}$
1	(-5.0, -1.0)	(-1.67, -1.0)	(1.67, -1.0)	(5.0, -1.0)
2	(-5.0, -0.34)	(-1.67, -0.34)	(1.67, -0.34)	(5.0, -0.34)
3	(-5.0, 0.34)	(-1.67, 0.34)	(1.67, 0.34)	(5.0, 0.34)
4	(-5.0, 1.0)	(-1.67, 1.0)	(1.67, 1.0)	(5.0, 1.0)

The knot vectors along  $u$ -, and  $v$ -directions are, respectively,

$$\begin{aligned} &0\ 0\ 0\ 0\ 1\ 1\ 1\ 1, \\ &0\ 0\ 0\ 0\ 1\ 1\ 1\ 1. \end{aligned}$$

## References

- [1] T.J.R. Hughes, J.A. Cottrell, and Y. Bazilevs. Isogeometric analysis: Cad, finite elements, nurbs, exact geometry and mesh refinement. *Computer methods in applied mechanics and engineering*, 194(39):4135–4195, 2005.
- [2] F. Auricchio, L. Beirão da Veiga, TJR Hughes, A. Reali, and G. Sangalli. Isogeometric collocation methods. *Mathematical Models and Methods in Applied Sciences*, 20(11):2075–2107, 2010.
- [3] Xiong Zhang, Xiao-Hu Liu, Kang-Zu Song, and Ming-Wan Lu. Least-squares collocation meshless method. *International Journal for Numerical Methods in Engineering*, 51(9):1089–1100, 2001.
- [4] Bo-Nan Jiang. Least-squares meshfree collocation method. *International Journal of Computational Methods*, 9(02), 2012.
- [5] Do Wan Kim and Yongsik Kim. Point collocation methods using the fast moving least-square reproducing kernel approximation. *International Journal for Numerical Methods in Engineering*, 56(10):1445–1464, 2003.
- [6] YJ Dai, XH Wu, and WQ Tao. Weighted least-squares collocation method (wlscm) for 2-d and 3-d heat conduction problems in irregular domains. *Numerical Heat Transfer, Part B: Fundamentals*, 59(6):473–494, 2011.
- [7] MH Afshar, M Lashckarbolok, and G Shobeyri. Collocated discrete least squares meshless (cdlsm) method for the solution of transient and steady-state hyperbolic problems. *International journal for numerical methods in fluids*, 60(10):1055–1078, 2009.
- [8] Bernard BT Kee, GR Liu, and C Lu. A least-square radial point collocation method for adaptive analysis in linear elasticity. *Engineering analysis with boundary elements*, 32(6):440–460, 2008.
- [9] L.A. Piegl and W. Tiller. *The NURBS book*. Springer Verlag, 1997.
- [10] C. De Boor. *A practical guide to splines*, volume 27. Springer Verlag, 2001.
- [11] Y. Zhang, W. Wang, and T.J.R. Hughes. Solid t-spline construction from boundary representations for genus-zero geometry. *Computer Methods in Applied Mechanics and Engineering*, 2012.
- [12] E. Cohen, T. Martin, RM Kirby, T. Lyche, and RF Riesenfeld. Analysis-aware modeling: Understanding quality considerations in modeling for isogeometric analysis. *Computer Methods in Applied Mechanics and Engineering*, 199(5):334–356, 2010.
- [13] Y. Bazilevs, VM Calo, JA Cottrell, JA Evans, TJR Hughes, S. Lipton, MA Scott, and TW Sederberg. Isogeometric analysis using t-splines. *Computer Methods in Applied Mechanics and Engineering*, 199(5):229–263, 2010.
- [14] M.R. Dörfel, B. Jüttler, and B. Simeon. Adaptive isogeometric analysis by local  $h$ -refinement with t-splines. *Computer methods in applied mechanics and engineering*, 199(5):264–275, 2010.
- [15] H.J. Kim, Y.D. Seo, and S.K. Youn. Isogeometric analysis for trimmed cad surfaces. *Computer Methods in Applied Mechanics and Engineering*, 198(37):2982–2995, 2009.
- [16] D. Burkhart, B. Hamann, and G. Umlauf. Iso-geometric finite element analysis based on catmull-clark subdivision solids. In *Computer Graphics Forum*, volume 29, pages 1575–1584. Wiley Online Library, 2010.
- [17] F. Auricchio, L. Beirão da Veiga, A. Buffa, C. Lovadina, A. Reali, and G. Sangalli. A fully locking-free isogeometric approach for plane linear elasticity problems: a stream function formulation. *Computer methods in applied mechanics and engineering*, 197(1):160–172, 2007.
- [18] T. Elguedj, Y. Bazilevs, VM Calo, and TJR Hughes.  $\bar{B}$  and  $\bar{F}$  projection methods for nearly incompressible linear and non-linear elasticity and plasticity using higher-order nurbs elements. *Comput. Methods Appl. Mech. Engrg.*, 197:2732–2762, 2008.
- [19] JA Cottrell, A. Reali, Y. Bazilevs, and TJR Hughes. Isogeometric analysis of structural vibrations. *Computer methods in applied mechanics and engineering*, 195(41):5257–5296, 2006.
- [20] T.J.R. Hughes, A. Reali, and G. Sangalli. Duality and unified analysis of discrete approximations in structural dynamics and wave propagation: Comparison of  $p$ -method finite elements with  $k$ -method nurbs. *Computer methods in applied mechanics and engineering*, 197(49):4104–4124, 2008.
- [21] W.A. Wall, M.A. Frenzel, and C. Cyron. Isogeometric structural shape optimization. *Computer Methods in Applied Mechanics and Engineering*, 197(33):2976–2988, 2008.
- [22] Y. Bazilevs, VM Calo, T.J.R. Hughes, and Y. Zhang. Isogeometric fluid-structure interaction: theory, algorithms, and computations. *Computational Mechanics*, 43(1):3–37, 2008.
- [23] Y. Bazilevs, VM Calo, Y. Zhang, and T.J.R. Hughes. Isogeometric fluid-structure interaction analysis with applications to arterial blood flow. *Computational Mechanics*, 38(4):310–322, 2006.
- [24] Y. Bazilevs, JR Gohean, TJR Hughes, RD Moser, and Y. Zhang. Patient-specific isogeometric fluid-structure interaction analysis of thoracic aortic blood flow due to implantation of the jarvik 2000 left ventricular assist device. *Computer Methods in Applied Mechanics and Engineering*, 198(45):3534–3550, 2009.
- [25] Y. Bazilevs, L. Beirão da Veiga, JA Cottrell, TJR Hughes, and G. Sangalli. Isogeometric analysis: approximation, stability and error estimates for  $h$ -refined meshes. *Mathematical Models and Methods in Applied Sciences*, 16(07):1031–1090, 2006.
- [26] JA Cottrell, TJR Hughes, and A. Reali. Studies of refinement and continuity in isogeometric structural analysis. *Computer methods in applied mechanics and engineering*, 196(41):4160–4183, 2007.



- [27] T.J.R. Hughes, A. Reali, and G. Sangalli. Efficient quadrature for nurbs-based isogeometric analysis. *Computer methods in applied mechanics and engineering*, 199(5):301–313, 2010.
- [28] M. Aigner, C. Heinrich, B. Jüttler, E. Pilgerstorfer, B. Simeon, and A. Vuong. Swept volume parameterization for isogeometric analysis. *Mathematics of Surfaces XIII*, pages 19–44, 2009.
- [29] G. Xu, B. Mourrain, R. Duvigneau, and A. Galligo. Optimal analysis-aware parameterization of computational domain in 3d isogeometric analysis. *Computer-Aided Design*, 45(4):812–821, 2013.
- [30] J.A. Cottrell, T.J.R. Hughes, and Y. Bazilevs. *Isogeometric analysis: toward integration of CAD and FEA*. Wiley, 2009.
- [31] Dominik Schillinger, John A Evans, Alessandro Reali, Michael A Scott, and Thomas JR Hughes. Isogeometric collocation: Cost comparison with galerkin methods and extension to adaptive hierarchical nurbs discretizations. *Computer Methods in Applied Mechanics and Engineering*, 267:170–232, 2013.
- [32] Hongwei Lin, Qianqian Hu, and Yunyang Xiong. Consistency and convergence properties of the isogeometric collocation method. *Computer Methods in Applied Mechanics and Engineering*, 267:471–486, 2013.
- [33] F Auricchio, L Beirão da Veiga, TJR Hughes, A Reali, and G Sangalli. Isogeometric collocation for elastostatics and explicit dynamics. *Computer methods in applied mechanics and engineering*, 249:2–14, 2012.
- [34] L Beirão da Veiga, C Lovadina, and A Reali. Avoiding shear locking for the timoshenko beam problem via isogeometric collocation methods. *Computer methods in applied mechanics and engineering*, 241:38–51, 2012.
- [35] F Auricchio, L Beirão da Veiga, J Kiendl, C Lovadina, and A Reali. Locking-free isogeometric collocation methods for spatial timoshenko rods. *Computer Methods in Applied Mechanics and Engineering*, 263(15):113–126, 2013.
- [36] Gene H Golub and Charles F van Van Loan. *Matrix computations (johns hopkins studies in mathematical sciences)*. 1996.
- [37] Pavel Solin. *Partial differential equations and the finite element method*. Wiley-Interscience, 2006.